# The Amino-acid Sequence in the Glycyl Chain of Insulin

## 1. THE IDENTIFICATION OF LOWER PEPTIDES FROM PARTIAL HYDROLYSATES

By F. SANGER* AND E. O. P. THOMPSON†
*Biochemical Laboratory, University of Cambridge*

(*Received* 16 *July* 1952)

Insulin has been shown to be composed of two types of polypeptide chains held together by —S—S— bridges. These chains may be separated after breaking the —S—S— bridges by oxidation to $—SO_3H$ groups (Sanger, 1949*a*). The more acidic glycyl chain is found in the fraction (*A*) which is soluble at pH 6·5 and the more basic phenylalanyl chain in the insoluble fraction (*B*). By making use of the methods of Consden, Gordon & Martin (1947) for fractionating peptides on paper chromatograms, Sanger & Tuppy (1951*a*, *b*) were able to determine the complete amino-acid sequence of the phenylalanyl chain. This and the following paper describe a similar investigation on fraction *A* to determine the amino-acid sequence in the glycyl chain.

End-group analysis of fraction *A* showed that it contained about twenty amino-acid residues (Sanger, 1949*a*) and this was in reasonable agreement with the value of 2900 for the molecular weight which was obtained in the ultracentrifuge (Gutfreund & Ogston, 1949). The amino-acid composition was considerably simpler than that of insulin or fraction *B* since it contained no arginine, histidine, lysine, phenylalanine, threonine or proline. By studying the dinitrophenyl (DNP) peptides produced on partial hydrolysis of the DNP derivative of fraction *A* it was shown that the *N*-terminal sequence was Gly.Ileu.Val.Glu.Glu (Sanger, 1949*b*).

The abbreviations used and the methods of writing peptide structures are the same as those given by Sanger & Tuppy (1951*a*).

## METHODS

### *Preparation of fraction A*

When the fractions of oxidized insulin were prepared as previously described (Sanger, 1949*a*) the tyrosine residues were to some extent oxidized to a product of unknown structure, which was distinguished from tyrosine by its $R_F$ value in butanol-acetic acid and was referred to as TyrX (Sanger & Tuppy, 1951*a*). This increased the complexity of the chromatograms since, for every tyrosine-containing peptide, two spots were present, one with tyrosine and one with TyrX. In the original method a large excess of performic acid was used to oxidize the cystine residues. It has now been found that if only a slight excess is used the formation of TyrX can be avoided. The yield of the fraction *A* is somewhat lower, but for most purposes this drawback is offset by the advantage of having a simpler peptide mixture.

Insulin (1 g.) was dissolved in 8 ml. formic acid and treated with 0·48 ml. 30% (w/w) $H_2O_2$ (1·6 equiv. on the basis of the S content). After standing for 1 hr. 8 ml. water were added to decompose the performic acid. Subsequent treatment and preparation of the fractions was as described before (Sanger, 1949*a*). The yield of fraction *A* was 0·16–0·22 g. and of fraction *B* 0·20–0·27 g. No TyrX could be detected in either fraction.

### *Hydrolysis and preliminary fractionation*

In the experiments described in this paper, hydrolysis was carried out in 12 N-HCl at 37°, the most generally useful conditions for obtaining a large number of small peptides. Hydrolysates prepared by the action of 0·1 N-HCl or N-NaOH at 100° were also studied, but they are not described as no new peptides were found in them, and they were not investigated in great detail.

To effect preliminary group fractionation of the hydrolysate prior to paper chromatography several methods have been studied in different experiments. An 'aromatic' fraction ($A_{1a}$) could be obtained by adsorption on charcoal. This method did not give a completely clear-cut separation since the fraction contained considerable quantities of the longer non-aromatic peptides besides those of tyrosine. However, it was sufficiently simple to study by paper chromatography and the method was useful for the initial removal of aromatic peptides, which do not fractionate well by ionophoresis in silica gel.

About half the residues in fraction *A* are acidic and there are no basic amino-acids. The majority of the peptides are therefore acidic and it was found that the most useful methods for group separations were those which fractionate the peptides according to the different number of charged groups they contain.

In a preliminary experiment the peptides were fractionated on the ion-exchange resin, Amberlite IR-4 B (Consden, Gordon & Martin, 1948; Sanger & Tuppy, 1951*a*). This separated the hydrolysate into three clear-cut fractions containing peptides of cysteic acid, other acidic peptides and neutral peptides, respectively. In general, however, ionophoretic methods were preferred since they are easier to carry out and separate the hydrolysate into a larger number of simpler fractions. Although better separations are usually obtained by ionophoresis on filter paper, it is not convenient to work with sufficient material on this support-

* Member of the scientific staff of the Medical Research Council.

† Present address: Wool Textile Research Laboratories, 343 Royal Parade, Parkville, Victoria, Australia.

ing medium. We have therefore used silica jelly as in the original method of Consden, Gordon & Martin (1946). Difficulties due to adsorption are encountered in this method with peptides containing basic or aromatic amino-acids; the former were absent from fraction *A*, the latter had been removed by adsorption on charcoal.

*Experiment* $A_1$. This is given as a typical experiment in which all the peptides detected are recorded. It is described as a single experiment, although much of the data was in fact obtained using other fractions from experiments in which slightly different methods for preliminary separations had been employed. Fraction *A* (100 mg.) was hydrolysed for 3 days at 37° with 5 ml. 12 N-HCl. After removal of the excess acid by repeated evaporation *in vacuo*, the residue was dissolved in 4 ml. 5% (w/v) acetic acid and stirred for 20 min. with 150 mg. charcoal (British Drug Houses Ltd., previously washed well with 20% (w/v) acetic acid). The charcoal was filtered off and washed with 2 ml. 5% acetic acid. The combined filtrates, containing the unadsorbed peptides, were taken to dryness and fractionated by ionophoresis as described below. The adsorbed peptides were eluted from the charcoal with 50 ml. 5% phenol in 20% acetic acid to give fraction $A_{1a}$. For the ionophoresis of the unadsorbed peptides the method (C 1) of Consden *et al.* (1946) was used. The jelly, which was strengthened with filter-paper pulp, was prepared by neutralizing 140 ml. 8% sodium silicate solution with 5% acetic acid (115 ml.). After setting it was washed overnight with 0·05 M-ammonium acetate to replace the $Na^+$ ions by $NH_4^+$ ions. The material to be fractionated was applied in a gutter near the cathode. The jelly was covered by a glass lid weighted to prevent uneven shrinking and puddling of the electrolyte. A potential of 220 V. was applied for 19 hr. At the end of this period a 'print' of the jelly was taken on to a sheet of filter paper, which was sprayed first with 1% methanolic KOH to remove $NH_3$, and then with 5% methanolic acetic acid followed by 0·1% ninhydrin in butanol. As the resolution was satisfactory, the jelly was cut out according to the marker sheet and the cuts dried *in vacuo* over $H_2SO_4$ and NaOH. The amino-acids and peptides were eluted from the dried jelly with 5% acetic acid. The solution was then filtered and evaporated to dryness. Fig. 1 is a diagram of the ionophoresis and shows the bands obtained. Each fraction was then investigated by paper chromatography.

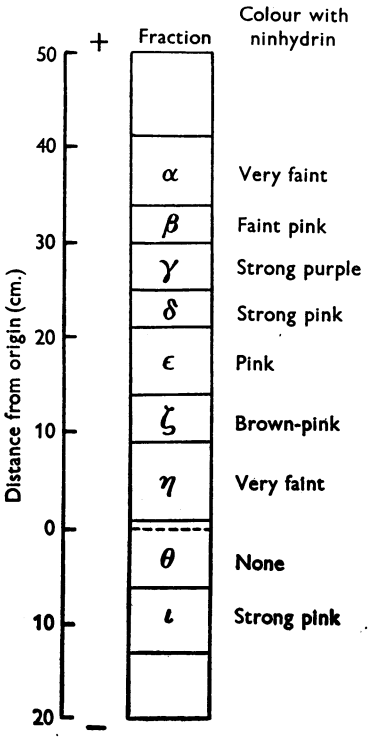


Fig. 1. Ionophoretic fractionation of unadsorbed peptides from partial hydrolysate of fraction *A* in 0·05 M-ammonium acetate (Exp. $A_1$).

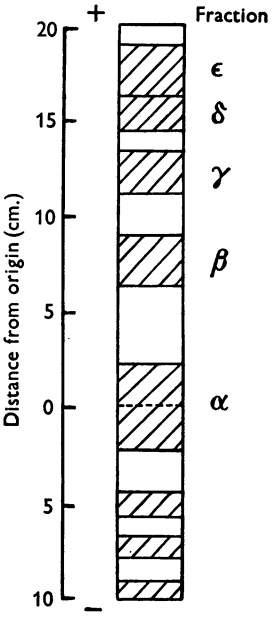


Fig. 2. Ionophoretic fractionation of cysteic acid peptides from partial hydrolysate of fraction *A* in 0·2 N-acetic acid (Exp. $A_2$).

*Experiment* $A_2$. Several peptides containing more than one cysteic acid residue were present in the hydrolysate of fraction *A*, and these could not easily be fractionated on paper chromatograms as they moved so slowly in all solvents studied. An attempt was therefore made to separate them by ionophoresis. Since the $-SO_3H$ group is more acidic (pK = approx. 1·3; Andrews & Schmidt, 1927) than the —COOH group (pK = 2·5) it is possible, by choosing an electrolyte of suitable pH, to find conditions where the latter are virtually uncharged whereas the former are ionized. Consden *et al.* (1946) made use of this property to separate cysteic acid from aspartic and glutamic acids with 0·02 N-HCl as electrolyte. We have found that good separations of cysteic acid peptides can be obtained by ionophoresis in 0·2 N-acetic acid (pH 2·75). Under these conditions all peptides with no $-SO_3H$ group move towards the cathode (due to the fully charged $-NH_3$ group); those with

one —$SO_3H$ group are approximately neutral and a very clear-cut fraction ($A_{2\alpha}$) with only this type of peptide can be obtained. Peptides with more than one cysteic acid residue move towards the anode, and in the case of those from fraction *A* can readily be fractionated. Free cysteic acid, which contains an unusually strong α-COOH group, also moves towards the anode.

Ionophoresis on paper was found to be very effective for this purpose, and the technique of Durrum (1950) was superior to several others that were investigated. This is probably due largely to the flow gradient of electrolyte up the paper, which tends to sharpen the bands. Since acetic acid is not itself an effective buffer it was important to remove free and bound acid from the hydrolysate and to apply the material at a relatively low concentration.

Fraction *A* was hydrolysed for 2 days with 12 N-HCl at 37°, and the excess acid removed by repeated evaporation *in vacuo*. The residue was taken up in water and excess $Ag_2O$ added to eliminate the bound HCl. After centrifuging and washing the precipitate with water, the solution was taken to dryness. It was applied in a concentration of 100 mg./ml. as a narrow band at the apex of the filter paper (Whatman no. 3, 56 × 25 cm.). After wetting the paper with electrolyte a potential of 220 V. was applied for 16 hr. The distribution of the bands, which was revealed by spraying a marker strip with ninhydrin, is shown in Fig. 2. They were cut out and eluted from the paper, taken to dryness and investigated as described in the next section.

For large 'cuts', such as those obtained in this experiment, a more convenient method of elution than that described by Sanger & Tuppy (1951*a*) was to allow the eluates to drop on to a sheet of celluloid or directly on to the polythene strip, thereby eliminating the need for constant watching during the elution.

### *Examination of the peptides*

The various fractions obtained in experiment $A_1$ were investigated by paper chromatography using essentially the same methods reported by Sanger & Tuppy (1951*a*). Usually Whatman no. 3 filter paper has been used since it has a high capacity and gives good resolutions. The solvent system phenol-0·3% $NH_3$, which was used extensively, will be referred to throughout simply as 'phenol', and *m*-cresol-0·3% $NH_3$ will be referred to as 'cresol'. With the exception of leucine and isoleucine, all the amino-acids present in fraction *A* (see Table 13) can be separated by chromatography using phenol. One-dimensional chromatograms were therefore used to identify the amino-acids present in the hydrolysates of the various peptides. Butanol-benzyl alcohol-HCN was used to distinguish between leucine and isoleucine.

To determine the *N*-terminal residues of the peptides, the DNP method has been used in most cases. Sanger & Tuppy (1951*a*) carried out the condensation between the peptide and 1:2:4-fluorodinitrobenzene (FDNB) in the presence of $NaHCO_3$. The salt had to be removed before paper chromatography could be carried out and this was achieved by extraction of the DNP peptide into ethyl acetate. In order to avoid this somewhat tedious step the reaction is now carried out in the presence of trimethylamine, which may be removed *in vacuo*. The identity of the *N*-terminal residue was deduced by Sanger & Tuppy from its absence in the hydrolysate of the DNP peptide. This clearly is less decisive than the identification of the actual DNP amino-acid present, and interpretation is especially difficult where tyrosine is present. In most cases we have therefore identified the DNP derivatives using buffered paper chromatograms (Blackburn & Lowther, 1951; Biserte & Osteaux, 1951; Monnier & Penasse, 1950). Sufficient material for such chromatograms could usually be obtained and, since the amino-acids present in the peptides were already known, it was generally possible to obtain an unequivocal result.

The peptide eluted from two chromatograms on Whatman no. 3 filter paper was dissolved in 0·1 ml. 1% trimethylamine and to this was added a solution of 0·01 ml. FDNB in 0·2 ml. ethanol. After standing for 2 hr., a few drops of water and trimethylamine solution were added and the excess FDNB extracted three times with ether. The residue, after evaporation of the aqueous solution to dryness, was taken up in three drops 5·7 N-HCl into a capillary tube. After 8 hr. heating at 105°, the yellow hydrolysate was diluted with twice its volume of water and the DNP amino-acid extracted into ether (three times). Both the ether extract and the aqueous solution were evaporated to dryness *in vacuo*. In general, the residue from the aqueous solution was chromatographed in phenol. DNP-cysteic acid and *O*-DNP-tyrosine are the only water-soluble DNP derivatives encountered with fraction *A*. DNP-cysteic acid has an $R_F$ value of 0·52. *O*-DNP-tyrosine travels at the solvent front and because of the dirty material collected there and its weak grey colour with ninhydrin, it is not easy to detect. Using butanol-acetic acid, both DNP-cysteic acid ($R_F = 0{\cdot}46$) and *O*-DNP-tyrosine ($R_F = 0{\cdot}84$) can be detected, but overlapping of the amino-acids, aspartic acid, serine, glycine, and possibly glutamic acid, then occurs.

The DNP amino-acids can mostly be resolved using *tert.*-amyl alcohol as solvent (Blackburn & Lowther, 1951). The paper (Whatman no. 1) and solvent were both treated with a phthalate buffer at pH 6. The chromatograms were preferably developed in the dark at a constant temperature (20°) (Biserte & Osteaux, 1951). The $R_F$ values obtained with this system were not constant and control spots of DNP amino-acids were always run simultaneously. Unless otherwise stated, the DNP amino-acids recorded in Tables 1–10 were identified using this solvent system.

In some of the earlier experiments, before the description of the above solvent system was published, use was made of the same solvent but a different buffer mixture (phosphate, pH 6·75) which had been developed by Partridge & Davis (unpublished). This did not give such good results as the phthalate buffer. Another solvent mixture suggested by Partridge & Davis was prepared from equal volumes of water, benzene and acetic acid, and could be used with unbuffered paper. It was useful for the DNP amino-acids with low $R_F$ values, such as DNP-cysteic acid, DNP-aspartic acid, DNP-glutamic acid and DNP-serine.

A disadvantage with paper chromatograms of DNP derivatives are the yellow artifacts which are also formed. The main ones are probably 2:4-dinitroaniline and 2:4-dinitrophenol. The dinitroaniline travels near the front and does not complicate the interpretation of results. The dinitrophenol spot is usually very strong and is accompanied by two other spots which travel at very similar rates to some of the DNP amino-acids. These artifact spots can be distinguished and removed, however, by exposure of the paper to HCl vapour, when they lose their yellow colour.

The identification of bis-DNP-tyrosine and some of the faster-moving DNP derivatives was somewhat unsatis-

Table 1. *Peptides from fraction* $A_{1a}$

| Spot no. (Figs. 3 and 4) | Amino-acids present | Strength of amino-acid after Hydrolysis | Strength of amino-acid after DNP treatment | DNP amino-acid identified | Structure* |
|---|---|---|---|---|---|
| 1 | $CySO_3H$ | × × | | | |
| | Glu | × | | | [$CySO_3H$, Glu, Ala] |
| | Ala | × | | | |
| 2 | $CySO_3H$ | × × × | | | |
| | Asp | × × × | | | [$CySO_3H$, Asp, Glu, Tyr] |
| | Glu | × × × | | | |
| | Tyr | × × | | | |
| 3 | $CySO_3H$ | × | | | |
| | Asp | × | | | [$CySO_3H$, Asp, Glu, Tyr] |
| | Glu | × | | | |
| | Tyr | × | | | |
| 4 | $CySO_3H$ | × × | × × | | |
| | Asp | × × | ? | Asp† | Asp.[$CySO_3H$, Tyr] |
| | Tyr | × × | ×‡ | | |
| 5 | $CySO_3H$ | × × × | × × | | |
| | Asp | × × × | × | Asp | Asp.[$CySO_3H$, Asp, Tyr] |
| | Tyr | × × × | ×‡ | | |
| 6 | $CySO_3H$ | × × | × | | |
| | Asp | × × | × | Asp | Asp.[$CySO_3H$, Asp, Tyr] |
| | Tyr | × × | ×‡ | | |
| 7 | $CySO_3H$ | × × | × × | | |
| | Tyr | × × | ?‡ | Tyr | Tyr.$CySO_3H$ |
| 8 | $CySO_3H$ | × × | | | |
| | Glu | × × | | | |
| | Gly | × | | Gly | Gly.[$CySO_3H$, Glu, Val, Leu] |
| | Val | × × | | | |
| | Leu | × × | | | |
| 9 | Asp | × | ? | | |
| | Tyr | × | ×‡ | Asp | Asp.Tyr |
| 10 | Asp | × × | ? | | |
| | Tyr | × × | ×‡ | Asp | Asp.Tyr |
| 11 | Asp | × × | × × | | |
| | Glu | × × | × | Glu | Glu.[Asp, Tyr] |
| | Tyr | × × | ×‡ | | |
| 12 | Glu | × × | × × | | |
| | Tyr | × × | ?‡ | Tyr | Tyr.Glu |
| 13 | Tyr | × × × | | | Tyrosine |
| 14 | Asp | × × | | | |
| | Tyr | × × | | | [Asp, Tyr] |
| 15 | Glu | × × × | | | |
| | Gly | × | | | |
| | Val | × × | | | [Glu, Gly, Val, Leu] |
| | Leu | × × | | | |
| 16 | Glu | × × × | | | |
| | Val | × × | | | [Glu, Val, Leu] |
| | Leu | × × | | | |
| 17 | Glu | × × | | | |
| | Tyr | × × | | | [Glu, Tyr, Leu] |
| | Leu | × × | | | |
| 18 | Glu | × | | | |
| | Gly | × | | | |
| | Val | × | | | Gly.[Glu, Val, Leu] ($A_{1\gamma 3}$) |
| | Leu | × | | | |

Table 1 (*cont.*)

| Spot no. (Figs. 3 and 4) | Amino-acids present | Strength of amino-acid after Hydrolysis | Strength of amino-acid after DNP treatment | DNP amino-acid identified | Structure* |
|---|---|---|---|---|---|
| 19 | Ser | × × × | × | | |
| | Tyr | × × × | ×‡ | Ser† | Ser.[Tyr, Leu] |
| | Leu | × × × | × × | | |
| 20 | Glu | × × | × | | |
| | Tyr | × × | ×‡ | Leu | Leu.[Glu, Tyr] |
| | Leu | × × | — | | |
| 21 | Tyr | × × | | | [Tyr, Leu] |
| | Leu | × × | | | |
| 22 | Glu | × | × | | |
| | Ser | × | ? | Ser | Ser.[Glu, Tyr, Leu] |
| | Tyr | × | § | | |
| | Leu | × | × | | |
| 23 | Glu | × × | | | |
| | Val | × × | | | Leu.[Glu, Val] ($A_{17,4}$) |
| | Leu | × × | | | |

* Square brackets indicate that the sequence of amino-acids inside the brackets has not been established.
† Identified using benzene-acetic acid-water.
‡ Identified as *O*-DNP-tyrosine.
§ *O*-DNP-tyrosine not identified on this phenol chromatogram.
Spots 8, 15, 18 and 19. Yellow colour with ninhydrin.
Spots 7 and 8 overlapped considerably, but as spot 8 gave an initial yellow colour with ninhydrin the separated areas of the spots could be detected and cut out. Better separated using collidine (see Fig. 13). After DNP treatment spot 7 gave a little DNP-glycine in addition to bis-DNP-tyrosine. Spots 11 and 12 often occurred as one spot and resolution could sometimes be achieved by attaching another sheet of filter paper to the bottom of the chromatogram and allowing the butanol-acetic acid to run for a longer time.
Spot 19. Yellow colour with ninhydrin. Sometimes difficult to locate as it was contaminated with red material washed down behind butanol-acetic acid front.
Spot 23 usually contained some tyrosine indicating contamination with some other peptide.

factory on the above solvent systems, since streaking often occurs near the front of the solvent and weak spots are difficult to detect. Better results were obtained with a reversed phase chromatogram using silicone-treated paper (Kritchevsky & Tiselius, 1951). A suitable solvent system was found to be $CHCl_3$ (10 vol.), *n*-propanol (10 vol.) and 1·5 N-acetic acid (6 vol.). After applying the spots to the paper, it was suspended in the tank with the organic phase (lower) and equilibrated for 3 hr. The paper was then dipped in the aqueous phase and the chromatogram allowed to develop. Bis-DNP-tyrosine and bis-DNP-lysine did not move, DNP-leucine ($R_F$=0·11), DNP-valine ($R_F$=0·25) and DNP-phenylalanine ($R_F$=0·13) moved slowly while all other DNP amino-acids and the 'artifact' bands ran faster. The $R_F$ values changed from experiment to experiment especially with the time of equilibration. The DNP amino-acids were therefore always identified by comparison with marker spots rather than by $R_F$ values.

In earlier experiments many of the peptides were subjected to deamination by the method of Consden *et al.* (1947), but these are not recorded since clearer results have since been obtained with the DNP method. Tyrosine was never detected after deamination, but if present in a peptide in a position other than *N*-terminal, two new spots were present, which moved on a phenol chromatogram slightly more slowly than serine and slightly faster than alanine, respectively. These spots, which are possibly nitroso derivatives of tyrosine, were pale yellow before treatment with ninhydrin and gave a greyish colour after treatment. If tyrosine was present as the *N*-terminal residue, these spots were absent or relatively weak and this fact could be used as a fairly reliable guide to the structure of aromatic peptides.

## RESULTS

### *Experiment* $A_1$

The results obtained with the various fractions of experiment $A_1$ are given in Figs. 3–12 and Tables 1–10. Figs. 3–12 are diagrams of the two-dimensional chromatograms. Tables 1–10 give the results obtained with the eluted peptides. In all cases column 2 lists the amino-acids given on complete hydrolysis and the approximate strength is indicated in column 3. (The × 's have the same significance as in the previous paper (Sanger & Tuppy, 1951*a*).) No distinction is made in these tables between leucine and isoleucine which are both referred to as Leu. Column 4 gives the strength of the amino-acid remaining in the aqueous solution after extraction of the hydrolysed DNP peptide solution with ether, and column 5 the DNP amino-acid identified in the ether extract or, in the case of DNP-cysteic acid, in the aqueous solution. The last column gives the structures of the peptides as far as they can be deduced from the data given in the tables.

Two chromatograms of fraction $A_{1a}$ are shown. In one (Fig. 3) the initial solvent was phenol, and in the other (Fig. 4) cresol. With the latter solvent better resolution of some of the faster-moving peptides was obtained. No attempt was made to identify the slower-moving spots.

Besides the peptides given in Table 1, fraction $A_{1a}$ contained appreciable amounts of Ser.Val.$CySO_3H$ ($A_{1\zeta_2}$), Leu.Glu.Leu ($A_{1\delta 6}$) and the free amino-acids cysteic acid, glycine and alanine. These were not present in earlier experiments where less charcoal (100 mg. charcoal/100 mg. fraction $A$ hydrolysate) was used to adsorb the aromatic peptides.

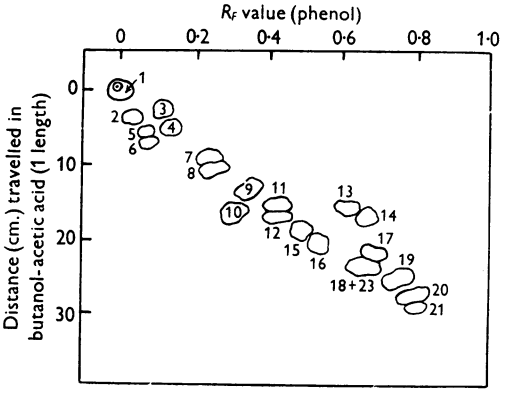


Fig. 3. Chromatogram of fraction $A_{1a}$ using phenol (see Table 1). In Figs. 3–14 the fraction was applied in the position marked by the small dotted circle, and run first in the direction represented horizontally, which was the longer dimension of the filter paper.

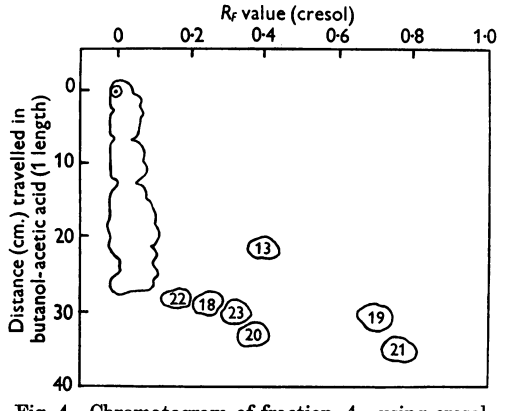


Fig. 4. Chromatogram of fraction $A_{1a}$ using cresol (see Table 1).

The duplication of many of the spots containing aspartic acid and tyrosine may be due to the presence of small amounts of TyrX, different amide contents, or due to different contents of aspartic acid residues. At least three spots (9, 10 and 14) were obtained which gave on hydrolysis aspartic acid and tyrosine and analysed as Asp.Tyr. This suggested that the aspartic acid residue concerned was derived initially from an asparagine residue and spot 14 may in fact be Asp(-$NH_2$).Tyr, as it gives a yellow colour with ninhydrin similar to asparagine. Again, spots 4–6 seem identical in amino-acid composition. The DNP treatment suggests that spots 5 and 6 differ from spot 4, in having a greater proportion of aspartic acid. Spots 1–6 were sometimes eluted together from a phenol chromatogram, which had been developed only one length in each direction, and the material re-run on a second chromatogram using phenol (2 lengths) and butanol-acetic acid (1½ or 2 lengths). Although much better resolutions were obtained in this way, there was never sufficient material available for a satisfactory end-group determination.

The chromatograms obtained with the ionophoretic fractions ($A_{1\alpha}$–$A_{1\iota}$) are much simpler than that obtained with the aromatic fraction ($A_{1a}$), only a few peptides being present in each. The results with fraction $A_{1\theta}$ are not recorded since it only contained traces of the monoamino-acids and Gly.[Glu, Val, Leu] ($A_{1\gamma 3}$).

*Partial hydrolysis of peptide* $A_{1a19}$. Peptide $A_{1a19}$ was further characterized by hydrolysing a sample with 12N-HCl for 3 days at 37°. The hydrolysate was analysed on a strip of Whatman no. 4 filter paper which was developed in only one direction with butanol-acetic acid (1 length). In addition to unchanged peptide and free amino-acids, a yellow spot ($R_F = 0{\cdot}65$) was obtained which yielded only serine and leucine on hydrolysis. It gave a yellow colour with ninhydrin, indicating that the serine was *N*-terminal and that it was therefore Ser.Leu (peptide $A_{1\iota 7}$). This established the structure of the original tripeptide ($A_{1a19}$) as Ser.Leu.Tyr. A spot very close to the unchanged tripeptide gave mainly leucine and tyrosine on hydrolysis but also contained some serine. It was probably Leu.Tyr mixed with a little Ser.Leu.Tyr.

*Partial hydrolysis of peptide* $A_{1\delta 6}$. This spot had a much lower $R_F$ value in phenol than Glu.Leu ($A_{1\iota 8}$) or Leu.Glu ($A_{1\iota 7}$). The strength of the amino-acids after DNP treatment suggested two glutamic residues were present. Samples of this peptide were collected and hydrolysed 3 days with 12N-HCl at 37°. The hydrolysate was chromatographed using phenol (1 length) and butanol-acetic acid (1 length) as solvents. Samples of Leu.Glu and Glu.Leu were run as markers during the development with butanol-acetic acid. The results are shown in Table 6. Spot 2 was unchanged tripeptide while spots 3 and 4 corresponded to Leu.Glu and Glu.Leu respectively. This establishes the structure of $A_{1\delta 6}$ as Glu.Leu.Glu since it has glutamic acid as the *N*-terminal residue and gives Glu.Leu and Leu.Glu on partial hydrolysis.

### *Experiment* $A_2$

Fraction $A_{2\alpha}$ (Fig. 2), which contained only peptides with one cysteic acid residue, was fractionated by chromatography and the results are shown in Fig. 13 and Table 11.

Samples of fraction $A_{2\beta}$ were chromatographed using phenol (2 lengths) followed by butanol-acetic acid (2 lengths) or collidine (1 length). In both cases

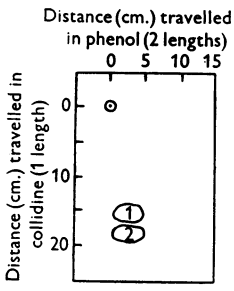


Fig. 5.

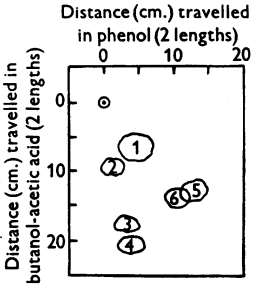


Fig. 6.

Fig. 5. Chromatogram of fraction $A_{1\alpha}$ (see Table 2). When the $R_F$ was low a second folded sheet was clamped to the lower edge of the original sheet and the distance of the spots from the original is then given instead of $R_F$.

Fig. 6. Chromatogram of fraction $A_{1\beta}$ (see Table 3).

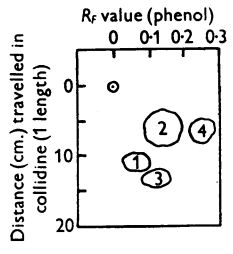


Fig. 7.

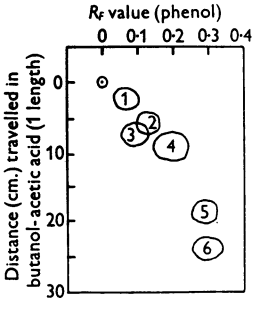


Fig. 8.

Fig. 7. Chromatogram of fraction $A_{1\gamma}$ (see Table 4).

Fig. 8. Chromatogram of fraction $A_{1\delta}$ (see Table 5).

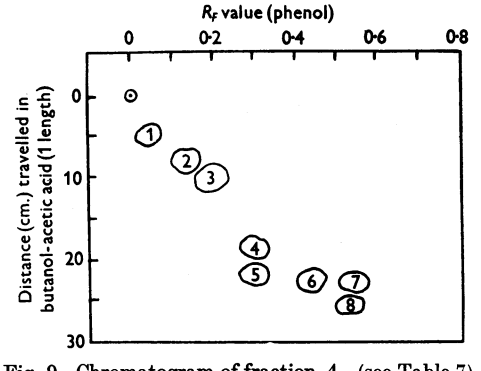


Fig. 9. Chromatogram of fraction $A_{1\epsilon}$ (see Table 7).

one main spot ($A_{2\beta 1}$) was obtained with smaller amounts of three other spots. The main spot gave a yellow colour with ninhydrin and had moved 9 cm. on phenol and 19 cm. on collidine or 12 cm. on phenol and 7·5 cm. on butanol-acetic acid. It contained $CySO_3H$ (× ×), Glu (× ×), Gly (×), Ala (× ×), Val (×), Leu (×).

Fraction $A_{2\gamma}$ was again submitted to iono-phoresis using the same conditions to remove some contamination from fraction $A_{2\delta}$. This gave a pure

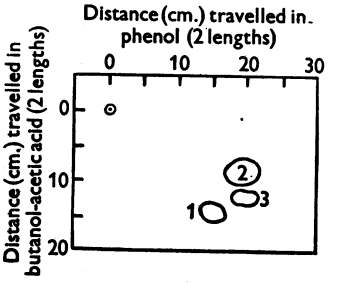


Fig. 10. Chromatogram of fraction $A_{1\zeta}$ (see Table 8).

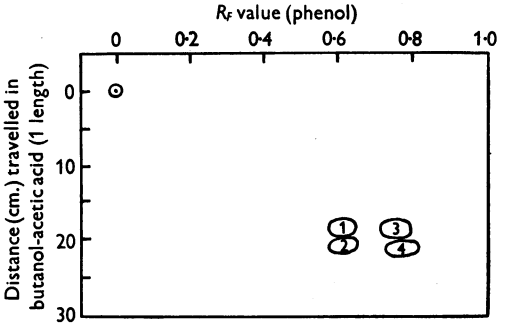


Fig. 11. Chromatogram of fraction $A_{1\eta}$ (see Table 9).

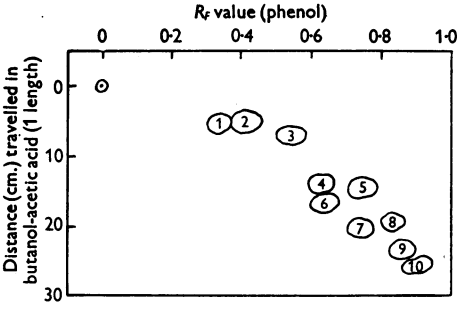


Fig. 12. Chromatogram of fraction $A_{1\iota}$ (see Table 10).

peptide ($A_{2\gamma 1}$) which contained $CySO_3H$ (× × ×), Glu (× ×), Ala (× ×). Glutamic acid was shown to be the *N*-terminal residue, indicating the structure Glu.[$CySO_3H$, $CySO_3H$, Ala].

A sample of this peptide was hydrolysed 3 days with 12 N-HCl at 37°, and the hydrolysate chromatographed on a sheet of Whatman no. 4 paper. The results are shown in Fig. 14 and Table 12. Glu.$CySO_3H$ was obtained so that the structure of $A_{2\gamma 1}$ is Glu.$CySO_3H$.[$CySO_3H$, Ala].

Table 2. *Peptides from fraction $A_{1\alpha}$*

| Spot no. (Fig. 5) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | $CySO_3H$ | × × × | × × × | Glu* | Glu.$CySO_3H$ |
| | Glu | × × × | × | | |
| 2 | $CySO_3H$ | × × × | × | $CySO_3H$† | $CySO_3H$.Asp |
| | Asp | × × × | × × × | | |

* Identified using benzene-water-acetic acid as solvent.
† Identified in aqueous layer using phenol.

Table 3. *Peptides from fraction $A_{1\beta}$*

| Spot no. (Fig. 6) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | $CySO_3H$ | × × × × | | | Cysteic acid |
| 2 | $CySO_3H$ | × | | | Glu.$CySO_3H$ ($A_{1\alpha 1}$) |
| | Glu | × | | | |
| 3 | Glu | × × × | × | Glu* | Glu.Asp |
| | Asp | × × × | × × × | | |
| 4 | Glu | × × × | | | See footnote |
| 5 | Asp | × × | | | Aspartic acid |
| 6 | $CySO_3H$ | × | | | $CySO_3H$.Ala ($A_{1\gamma 3}$) |
| | Ala | × | | | |

* Identified using benzene-water-acetic acid as solvent.
Spot 1. Spread over extensive area, due to salt or overloading.
Spot 4. Contained only glutamic acid but moves at a very different rate from free glutamic acid or glutamine. It is presumably Glu.Glu known to be present in the *N*-terminal sequence. Its rate, relative to cysteic acid, was similar to that found for the synthetic material by Consden, Gordon & Martin (1949), and the ionophoretic rate was faster than that of glutamic acid, in agreement with the results of these workers. On paper ionophoresis using 0·05 M-ammonium acetate, glutamic acid travelled 6·1 cm. while spot 4 travelled 7·7 cm.

Table 4. *Peptides from fraction $A_{1\gamma}$*

| Spot no. (Fig. 7) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | $CySO_3H$ | × × | | | Cysteic acid |
| 2 | Asp | × × × | | | Aspartic acid |
| 3 | $CySO_3H$ | × × × | × | $CySO_3H$* | $CySO_3H$.Ala |
| | Ala | × × × | × × × | | |
| 4 | Glu | × × | | | Glutamic acid |

* Identified in aqueous phase using butanol-acetic acid.
Spots 2 and 3 were not well separated when butanol-acetic acid was used as the second solvent.
Spot 3. Yellow colour with ninhydrin turning to grey on longer heating.

Table 5. *Peptides from fraction $A_{1\delta}$*

| Spot no. (Fig. 8) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | $CySO_3H$ | × × | | | Cysteic acid |
| 2 | Asp | × × | | | Aspartic acid |
| 3 | $CySO_3H$ | × × | | | $CySO_3H$.Ala ($A_{1\gamma 3}$) |
| | Ala | × × | | | |
| 4 | Glu | × × × × | | | Glutamic acid |
| 5 | Asp | × × | × × | Leu | Leu.[Asp, Glu] |
| | Glu | × × | × | | |
| | Leu | × × | — | | |
| 6 | Glu | × × × × | × × | Glu | Glu.[Glu, Leu] |
| | Leu | × × × | × × | | |

Table 6. *Partial hydrolysate of* $A_{1\delta 6}$

| Spot no. | $R_F$ in phenol | Distance moved in butanol-acetic acid (cm.) | Amino-acids present | Strength |
|---|---|---|---|---|
| 1 | 0·18 | 7·3 | Glu | × |
| 2 | 0·26 | 18·2 | Glu | × × |
| | | | Leu | × |
| 3 | 0·53 | 19 | Glu | × × |
| | | | Leu | × × |
| 4 | 0·53 | 21 | Glu | × |
| | | | Leu | × |
| 5 | 0·78 | 20 | Leu | × |

All leucine spots from the chromatogram of these peptide hydrolysates were cut out, washed with acetone to remove excess ninhydrin, eluted and the material rechromatographed using butanol-benzyl alcohol as solvent. Only leucine, and no isoleucine, was present.

Table 7. *Peptides from fraction* $A_{1\epsilon}$

| Spot no. (Fig. 9) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | $CySO_3H$ | × | | | Cysteic acid |
| 2 | $CySO_3H$ | × | | | $CySO_3H$.Ala ($A_{1\gamma 3}$) + aspartic acid |
| | Asp | × | | | |
| | Ala | × | | | |
| 3 | $CySO_3H$ | × | | | [$CySO_3H$, Val] ($A_{1\zeta 3}$) + glutamic acid |
| | Glu | × × | | | |
| | Val | × | | | |
| 4 | Asp | × × | | | Leu.[Asp, Glu] ($A_{1\delta 5}$) |
| | Glu | × × × | | | |
| | Leu | × × | | | |
| 5 | Glu | × × × | | | Glu.[Glu, Leu] ($A_{1\delta 6}$) |
| | Leu | × × | | | |
| 6 | Glu | × × × | | | Val.Glu + $A_{1a16}$ (see footnote) |
| | Val | × × | | | |
| | Leu | × | | | |
| 7 | Glu | × × × | × × × | Leu | Leu.Glu |
| | Leu | × × × | × | | |
| 8 | Glu | × × × | × | Glu | Glu.Leu |
| | Leu | × × × | × × × | | |

Spot 6 is probably a mixture of peptide $A_{1a16}$ ([Glu, Val, Leu]) with Val.Glu. In another experiment, in which more charcoal (approx. 1 g./100 mg. hydrolysate) had been used to remove the 'aromatic' fraction, Val.Glu was identified as a single spot in this position on a chromatogram from a similar ionophoretic fraction. Deamination with nitrosyl chloride indicated valine was *N*-terminal.

Table 8. *Peptides from fraction* $A_{1\zeta}$

| Spot no. (Fig. 10) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | Glu | × | | | Glutamic acid |
| 2 | $CySO_3H$ | × × × | × × × | Ser | Ser.[$CySO_3H$, Val] |
| | Ser | × × × | × | | |
| | Val | × × × | × × × | | |
| 3 | $CySO_3H$ | × | | | [$CySO_3H$, Val] |
| | Val | × | | | |

Besides the peptides given above, fraction $A_{1\zeta}$ also contained traces of Leu.Glu ($A_{1\epsilon 7}$) and Glu.Leu ($A_{1\epsilon 8}$) which were detected on a chromatogram in which each solvent was run only one length.

Spot 2. Yellow colour with ninhydrin.

Spot 3 gave a very faint colour with ninhydrin, and could more easily be detected by its fluorescence in ultraviolet light. Deamination with nitrosyl chloride indicated that valine was probably the *N*-terminal residue.

Table 9. *Peptides from fraction* $A_{1\eta}$

| Spot no. (Fig. 11) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | Structure |
|---|---|---|---|---|
| 1 | Glu | × | | Leu.Glu ($A_{1\epsilon 7}$) |
| | Leu | × | | |
| 2 | Glu | × | | Glu.Leu ($A_{1\epsilon 8}$) |
| | Leu | × | | |
| 3 | Glu | × | × | Gly.[Glu, Val, Leu] |
| | Gly | × | — | |
| | Val | × | × | |
| | Leu | × | × | |
| 4 | Glu | × × | × × | Leu.[Glu, Val] |
| | Val | × × | × × | |
| | Leu | × × | ? | |

Spot 3. Yellow colour with ninhydrin.

Table 10. *Peptides from fraction* $A_{1\iota}$

| Spot no. (Fig. 12) | Amino-acids present | Strength of amino-acid after: Hydrolysis | Strength of amino-acid after: DNP treatment | DNP amino-acid identified | Structure |
|---|---|---|---|---|---|
| 1 | Ser | × × | | | Serine |
| 2 | Gly | × × | | | Glycine |
| 3 | Ala | × × × | | | Alanine |
| 4 | Tyr | × × | | | Tyrosine |
| 5 | Val | × × | | | Valine |
| 6 | Ser | × × | ? | Ser | Ser.Val |
| | Val | × × | × × | | |
| 7 | Ser | × × | × | Ser | Ser.Leu |
| | Leu | × × | × × × | | |
| 8 | Leu | × × | | | Leucine |
| 9 | Ser | × | | | Ser.Leu.Tyr ($A_{1a19}$) |
| | Tyr | × | | | |
| | Leu | × | | | |
| 10 | Leu | × × | ? | Leu | Ileu.Val |
| | Val | × × | × × | | |

Spots 6, 7 and 9. Yellow colour with ninhydrin.
Spot 7 was shown to contain leucine and not isoleucine using butanol-benzyl alcohol.

Table 11. *Peptides from fraction* $A_{2\alpha}$

| Spot. no. (Fig. 13) | Strength | Structure | Identical with |
|---|---|---|---|
| 1 | × × | Glu.$CySO_3H$ | $A_{1\alpha 1}$ |
| 2 | × × | $CySO_3H$.Asp | $A_{1\alpha 2}$ |
| 3 | × | [$CySO_3H$, Asp, Glu, Tyr] | $A_{1a2}$, $A_{1a3}$ |
| 4 | × × | [$CySO_3H$, Asp, Tyr] | $A_{1a4}$, $A_{1a5}$ |
| 5 | × × × | $CySO_3H$.Ala | $A_{1\gamma 3}$ |
| 6 | × × | Tyr.$CySO_3H$ | $A_{1a7}$ |
| 7 | × | Gly.[$CySO_3H$, Glu, Val, Leu] | $A_{1a8}$ |
| 8 | × | [$CySO_3H$, Glu, Val, Leu] | — |
| 9 | × × × × | Ser.[$CySO_3H$, Val] | $A_{1\zeta 2}$ |

Spots 7 and 8. The glutamic acid spot was stronger than the other amino-acid spots.
Spots 7 and 9. Yellow colours with ninhydrin.
Spot 8. Blue colour with ninhydrin.

Fraction $A_{2\delta}$ was chromatographed using phenol (2 lengths) and collidine (1 length). Two spots were obtained. One ($A_{2\delta 2}$) was free cysteic acid. It had travelled 6·5 cm. in phenol and 16·5 cm. in collidine. The other ($A_{2\delta 1}$) had travelled 1·4 cm. in phenol and 20·7 cm. in collidine. It contained cysteic acid and alanine, and hydrolysis of the DNP derivative gave DNP-$CySO_3H$ (× ×), $CySO_3H$ (× ×), Ala (× ×). Its structure was therefore $CySO_3H$.[$CySO_3H$, Ala].

Fraction $A_{2\epsilon}$ gave only free cysteic acid on hydrolysis.

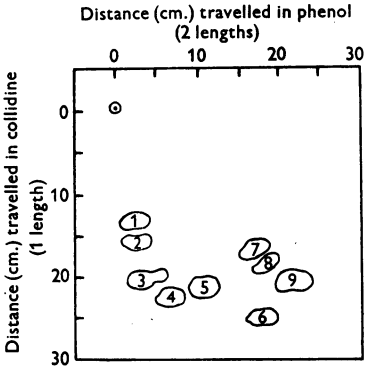


Fig. 13. Chromatogram of fraction $A_{2\alpha}$ (see Table 11).

## DISCUSSION

Since the amino-acid content of fraction *B* is accurately known from the structure determined by Sanger & Tuppy (1951*b*) it is possible to calculate the composition of fraction *A* as the difference between the analytical values for insulin and those for fraction *B*. Using the values collected by Tristram (1949) for the composition of insulin, the figures shown in column 2 of Table 13 are obtained for fraction *A*. In the third column are shown the values obtained in an approximate analysis done by paper chromatography (Sanger, 1949*a*).

### *Peptides of tyrosine*

The various tyrosine peptides are all found in fraction $A_{1a}$. There are four dipeptides containing tyrosine: Tyr.Glu ($A_{1a12}$), Asp.Tyr ($A_{1a9}$, $A_{1a10}$, $A_{1a14}$), Tyr.$CySO_3H$ ($A_{1a7}$) and [Tyr, Leu] ($A_{1a21}$). Since there are only two tyrosine residues in fraction *A* the structure of $A_{1a21}$ must be Leu.Tyr. This is confirmed by the structure of $A_{1a19}$ which was shown by partial hydrolysis to be Ser.Leu.Tyr. Peptide $A_{1a20}$ (Leu.[Glu, Tyr]) cannot be Leu.Glu.-Tyr since Glu.Tyr has not been identified, whereas Tyr.Glu ($A_{1a12}$) has. It must therefore be Leu.Tyr.-Glu and the sequence in fraction *A* becomes Ser.Leu.Tyr.Glu. Peptide $A_{1a22}$ presumably has this structure.

A DNP peptide containing those four amino-acids was detected as a contaminant of a DNP-glycyl peptide in partial hydrolysates of the DNP derivative of fraction *A* (Woolley, 1948; Sanger, 1949*b*). It presumably had the structure Ser.Leu.-(*O*-DNP) Tyr.Glu.

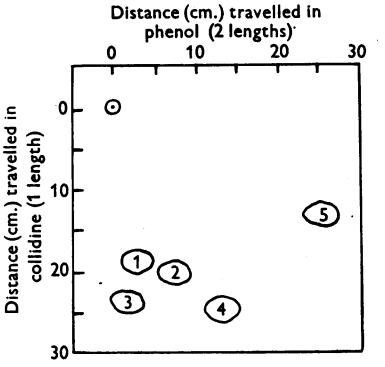


Fig. 14. Chromatogram of partial hydrolysate of peptide $A_{2\gamma 1}$ (see Table 12).

Table 12. *Partial hydrolysis of peptide* $A_{2\gamma 1}$

| Spot. no. (Fig. 14) | Strength | Structure | Identical with |
|---|---|---|---|
| 1 | × × | Glu.$CySO_3H$ | $A_{1a1}$ |
| 2 | × | Cysteic acid | — |
| 3 | × × | $CySO_3H$.[$CySO_3H$, Ala] | $A_{2\delta 1}$ |
| 4 | × × | $CySO_3H$.Ala | $A_{1\gamma 3}$ |
| 5 | × | Glutamic acid | — |

Table 13. *Probable amino-acid composition of fraction A*

| Amino-acid | No. of residues/mol. Calculated from insulin – fraction *B* | Found by approx. analysis |
|---|---|---|
| Cysteic acid | 4 | 4 |
| Aspartic acid | 2 | 2 |
| Glutamic acid | 4* | 4 |
| Serine | 2 | 2 |
| Glycine | 0–1 | 1 |
| Alanine | 1 | 1 |
| Tyrosine | 2 | 2 |
| Valine | 1 | 2 |
| Leucine | 2 | 3 |
| Isoleucine | 1–2 | 1 |

* The figures given by Tristram would indicate a value of 4·5 for the glutamic acid content. Recent analyses (Chibnall & Rees, 1951) make it almost certain that 4 is the correct figure.

Two other dipeptides (Asp.Tyr and Tyr.$CySO_3H$) are not included in the above sequence and must therefore contain the other tyrosine residue. The sequence is therefore Asp.Tyr.$CySO_3H$ and peptide $A_{1a4}$ (Asp.[$CySO_3H$, Tyr]) clearly has this structure. Peptide $A_{1a11}$ has an *N*-terminal glutamic acid residue and also contains aspartic acid and tyrosine.

Its structure is therefore Glu.Asp.Tyr and the sequence in fraction $A$ is Glu.Asp.Tyr.$CySO_3H$. This is confirmed by the presence of Glu.Asp ($A_{1\beta 3}$). The presence of free aspartic acid in the hydrolysate of the DNP derivatives of peptides $A_{1a5}$ and $A_{1a6}$ was taken as evidence that there was a second aspartic acid in these peptides. This would make the above sequence Glu.Asp.Tyr.$CySO_3H$.-Asp. However, the evidence is too slight to be certain of this.

### *N-terminal sequence*

Using the DNP technique it was shown that the $N$-terminal sequence was Gly.Ileu.Val.Glu.Glu (Sanger, 1949*b*). Confirmation of this structure is obtained from a number of peptides which are listed in Table 14. Owing to the stability of bonds involving isoleucine and valine many large peptides from this sequence are present, and are mostly found in the fraction adsorbed on charcoal ($A_{1a}$). The two peptides containing one glutamic acid residue (Ileu.Val.Glu and Gly.Ileu.Val.Glu) move rather rapidly on the phenol chromatogram in spot $A_{1a18}$. The two slower-moving spots ($A_{1a16}$, $A_{1a15}$) containing these same amino-acids presumably have two glutamic acid residues.

Since there is only one glycine residue in fraction $A$, any peptide containing glycine must be derived from this $N$-terminal sequence. Such a peptide is $A_{1a8}$. It contains the amino-acids of this sequence and also cysteic acid. Since it moves relatively slowly on a chromatogram, and is found in fraction $A_{2\alpha}$ (spot $A_{2\alpha 7}$) on ionophoresis in acetic acid, the cysteic acid must be an integral part of the peptide, whose structure is therefore Gly.Ileu.Val.Glu.-Glu.$CySO_3H$. The presence of Glu.$CySO_3H$ ($A_{1\alpha 1}$) confirms this sequence and peptide $A_{2\alpha 8}$ is presumably Ileu.Val.Glu.Glu.$CySO_3H$. Although the results hitherto discussed do not preclude the presence of another residue of leucine, valine or glutamic acid between the second glutamic acid and the cysteic acid residue in this sequence, it will be seen below that this is in fact not possible.

### *Peptides of alanine*

There is only one alanine residue in fraction $A$ and only one dipeptide was found. This was $CySO_3H$.Ala ($A_{1\gamma 3}$). Alanine was also found in peptides that moved towards the anode in Exp. $A_2$, that is to say in peptides containing two cysteic acid residues. Peptide $A_{2\delta 1}$ was $CySO_3H$.[$CySO_3H$, Ala] and peptide $A_{2\gamma 1}$ Glu.$CySO_3H$.[$CySO_3H$, Ala]. Clearly two sequences are possible Glu.$CySO_3H$.-$CySO_3H$.Ala or Glu.$CySO_3H$.Ala.$CySO_3H$. If the second structure were correct Ala.$CySO_3H$ should be detected in the hydrolysate. Consden & Gordon (1950) have shown it to be stable, for it was isolated from a 10-day partial hydrolysate of wool. They also showed that it could be separated from $CySO_3H$.Ala by ionophoresis or chromatography with collidine. In spite of many attempts, no evidence for the presence of this peptide in the fraction $A$ hydrolysate could be obtained and it is therefore concluded that the former sequence (Glu.$CySO_3H$.$CySO_3H$.Ala) is correct. Similarly, Glu.$CySO_3H$.Ala, which would be expected from the second sequence, was never detected. Unfortunately it has not been possible to obtain direct evidence for the presence of $CySO_3H$.$CySO_3H$. It would be expected to move slowly on chromatograms and is probably present in spot $A_{1a1}$ with $CySO_3H$.$CySO_3H$.Ala and Glu.$CySO_3H$.$CySO_3H$.-Ala. On ionophoresis in acetic acid it is probably in fraction $A_{2c}$ with cysteic acid.

The fact that Glu.$CySO_3H$ is present in the $N$-terminal position of the above sequence (Glu.-$CySO_3H$.$CySO_3H$.Ala) and in the $C$-terminal position in the sequence Gly.Ileu.Val.Glu.Glu.-$CySO_3H$ suggests that the two are joined together. This is confirmed from the composition of peptide $A_{2\beta 1}$, which contained the amino-acids of the $N$-terminal sequence and also alanine. Since it moved towards the anode on ionophoresis in acetic acid, it must contain the two adjacent cysteic acid residues and its structure is therefore Gly.Ileu.Val.Glu.-Glu.$CySO_3H$.$CySO_3H$.Ala.

### *Other peptides*

Three sequences have so far been deduced:

(1) Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.
(2) Ser.Leu.Tyr.Glu.
(3) Glu.Asp.Tyr.$CySO_3H$.

Apart from the valine peptides derived from sequence 1 there are three peptides containing valine: Ser.Val ($A_{1\iota 6}$), Val.$CySO_3H$ ($A_{1\zeta 3}$) and Ser.[$CySO_3H$, Val] ($A_{1\zeta 2}$). Since there are certainly not more than two valine residues in fraction $A$ these must all come from the same sequence: Ser.Val.$CySO_3H$. This structure may also be derived from considering the peptides of serine.

Aspartic acid is present in sequence 3 and also in the peptides $A_{1\delta 5}$ (Leu.[Asp, Glu]) and $A_{1\alpha 2}$ ($CySO_3H$.Asp). Since there are only two aspartic acid residues in fraction $A$ and since the above two peptides cannot come from the same sequence, peptide $A_{1\delta 5}$ must be joined to the $N$-terminal end of sequence 3, which thus becomes Leu.Glu.Asp.-Tyr.$CySO_3H$. Leu.Glu is found in peptide $A_{1\epsilon 7}$.

Peptide $A_{1\delta 6}$ was shown by partial hydrolysis to be Glu.Leu.Glu. Since there are only four glutamic acid residues this must be joined on to sequence 3 to give Glu.Leu.Glu.Asp.Tyr.$CySO_3H$. Any other way of fitting the sequences together would require at least five glutamic acid residues.

Table 14. *Peptides obtained in partial hydrolysate of fraction A of insulin*

| | | | | | |
|---|---|---|---|---|---|
| Dipeptides | Ileu.Val ($A_{1\iota 10}$) | Glu.$CySO_3H$ ($A_{1\alpha 1}$) | Ser.Leu ($A_{1\iota 7}$) | Glu.Leu ($A_{1\epsilon 8}$) | Asp.Tyr ($A_{1a9}$) | Ser.Val ($A_{1\iota 6}$) | $CySO_3H$.Asp ($A_{1\alpha 2}$) |
| | Val.Glu ($A_{1\epsilon 6}$) | | Leu.Tyr ($A_{1a21}$) | Leu.Glu ($A_{1\epsilon 7}$) | Tyr.$CySO_3H$ ($A_{1a7}$) | Val.$CySO_3H$ ($A_{1\zeta 3}$) | |
| | Glu.Glu ($A_{1\beta 4}$) | $CySO_3H$.Ala ($A_{1\gamma 3}$) | Tyr.Glu ($A_{1a12}$) | Glu.Asp ($A_{1\beta 3}$) | | | |
| Higher peptides | Ileu.Val.Glu ($A_{1\gamma 4}$) | $CySO_3H$.$CySO_3H$.Ala ($A_{2\delta 1}$) | Ser.Leu.Tyr ($A_{1a19}$) | Leu.Glu.Asp ($A_{1\delta 5}$) | | Ser.Val.$CySO_3H$ ($A_{1\zeta 2}$) | |
| | Gly.Ileu.Val.Glu ($A_{1\gamma 3}$) | | Leu.Tyr.Glu ($A_{1a20}$) | Glu.Asp.Tyr ($A_{1a11}$) | | | |
| | Glu.$CySO_3H$.$CySO_3H$.Ala ($A_{2\gamma 1}$) | | | | | | |
| | Ileu.Val.Glu.Glu ($A_{1a16}$) | | | | Asp.Tyr.$CySO_3H$ ($A_{1a4}$) | | |
| | Gly.Ileu.Val.Glu.Glu ($A_{1a15}$) | | | Glu.Leu.Glu ($A_{1\delta 6}$) | | | |
| | Ileu.Val.Glu.Glu.$CySO_3H$ ($A_{2\alpha 8}$) | | Ser.Leu.Tyr.Glu ($A_{1a22}$) | | Glu.Asp.Tyr.$CySO_3H$ ($A_{1a3}$) | | |
| | Gly.Ileu.Val.Glu.Glu.$CySO_3H$ ($A_{1a8}$, $A_{2\alpha 7}$) | | | | | | |
| | Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala ($A_{2\beta 1}$) | | | | | | |
| Sequences in fraction *A* | Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala | | Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$ | | | Ser.Val.$CySO_3H$ | $CySO_3H$.Asp |

In the three sequences deduced there are still five glutamic acid residues so that one must be common to two sequences. This can only be the *C*-terminal residue of sequence 2 and the *N*-terminal residue of sequence 3, which are therefore joined together, and the structure in fraction *A* becomes Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$. Both the 'Leu' residues in this sequence are leucine and not isoleucine since peptides $A_{1\iota 7}$ (Ser.Leu) and $A_{1\delta 6}$ (Glu.Leu.Glu) contain no isoleucine. The presence of peptide $A_{1a17}$ ([Glu, Tyr, Leu]), which was different from $A_{1a20}$ (Leu.Tyr.Glu), can readily be understood from this sequence. A number of structures are possible, such as Tyr.Glu.Leu or Leu.Tyr.Glu.Leu.Glu.

All the peptides encountered in this work are now accounted for in terms of the four sequences:

(1) Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.
(2) Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.
(3) Ser.Val.$CySO_3H$.
(4) $CySO_3H$.Asp.

All the amino-acids known to be present in fraction *A* are present in these sequences and there is one too many cysteic acid residues. The $CySO_3H$.Asp must therefore be joined to the *C*-terminal end of sequence 2 or 3, but it is not possible to decide conclusively which.

In conclusion, the various peptides identified are listed in Table 14 with the sequences deduced from them.

## SUMMARY

1. Fraction *A* of oxidized insulin has been subjected to partial hydrolysis with acid and the resulting peptides fractionated by charcoal adsorption, ionophoresis and paper chromatography.

2. It is concluded that the following sequences are present in this fraction:

Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala,
Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$,
Ser.Val.$CySO_3H$ and $CySO_3H$.Asp.

One of us (E. O. P. T.) wishes to thank the Commonwealth Government of Australia for a Commonwealth Scientific and Industrial Research Organization studentship. Part of the work described in this paper was carried out during the tenure by F.S. of a Beit Memorial Fellowship, which is gratefully acknowledged.

## REFERENCES

Andrews, S. & Schmidt, C. L. A. (1927). *J. biol. Chem.* **73**, 651.
Biserte, G. & Osteaux, R. (1951). *Bull. Soc. Chim. biol., Paris*, **33**, 50.
Blackburn, S. & Lowther, A. G. (1951). *Biochem. J.* **48**, 126.
Chibnall, A. C. & Rees, M. W. (1951). *Biochem. J.* **48**, xlvii.
Consden, R. & Gordon, A. H. (1950). *Biochem. J.* **46**, 8.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1946). *Biochem. J.* **40**, 33.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1947). *Biochem. J.* **41**, 590.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1948). *Biochem. J.* **42**, 443.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1949). *Biochem. J.* **44**, 548.
Durrum, E. L. (1950). *J. Amer. chem. Soc.* **72**, 2943.
Gutfreund, H. & Ogston, A. G. (1949). *Biochem. J.* **44**, 163.
Kritchevsky, T. H. & Tiselius, A. (1951). *Science*, **114**, 299,
Monnier, R. & Penasse, L. (1950). *C.R. Acad. Sci., Paris*, **230**, 1176.
Sanger, F. (1949*a*). *Biochem. J.* **44**, 126.
Sanger, F. (1949*b*). *Biochem. J.* **45**, 563.
Sanger, F. & Tuppy, H. (1951*a*). *Biochem. J.* **49**, 463.
Sanger, F. & Tuppy, H. (1951*b*). *Biochem. J.* **49**, 481.
Tristram, G. R. (1949). *Advanc. Protein Chem.* **5**, 83.
Woolley, D. W. (1948). *Fed. Proc.* **7**, 200.

# The Amino-acid Sequence in the Glycyl Chain of Insulin

## 2. THE INVESTIGATION OF PEPTIDES FROM ENZYMIC HYDROLYSATES

BY F. SANGER* AND E. O. P. THOMPSON†
*Biochemical Laboratory, University of Cambridge*

(*Received* 16 *July* 1952)

In the previous paper (Sanger & Thompson, 1953*a*) four sequences were shown to be present in fraction *A* of oxidized insulin. It was found, as in the case of fraction *B* (Sanger & Tuppy, 1951*a*), that the complete sequence could not be derived from the small peptides produced on acid hydrolysis, due largely to the great lability of the bonds involving the amino groups of the serine residues. It was therefore necessary to investigate the action of proteolytic enzymes, and the present paper describes the results obtained with peptic and chymotryptic hydrolysates. Trypsin was found to be without action on fraction *A*.

* Member of the scientific staff of the Medical Research Council.
† Present address: Wool Textile Research Laboratories, 343 Royal Parade, Parkville, Victoria, Australia.